American Journal of Medical Genetics (Neuropsychiatric Genetics) 74:289–295 (1997)

Cladistic Analysis of Disease Association With Tyrosine Hydroxylase: Application to Manic-Depressive Disease and Alcoholism

Elizabeth A. Lobos and Richard D. Todd*
Departments of Psychiatry and Genetics, Washington University School of Medicine, St. Louis, Missouri

We evaluated the involvement of tyrosine hydroxylase (TH) mutations in susceptibility to manic-depressive disease (MDD) and alcoholism (ALC) with a cladistics-based association analysis. Eighty-one probands with MDD, 113 probands with alcoholism, and 80 normal controls were tested for differences in frequency of nine haplotypes at the TH locus. The haplotypes were comprised of four restriction fragment length polymorphisms spanning the TH gene. A cladogram constructed from the haplotypes provided the evolutionary context for a nested statistical analysis. Statistically significant evidence was found for association of a subgroup of the sample for each of the disorders with different branches of the gene tree, but the findings were sensitive to minor changes in estimated haplotype frequencies. Am. J. Med. Genet. 74:289–295, 1997. © 1997 Wiley-Liss, Inc.

KEY WORDS: genetics; bipolar disorder; cladogram; gene tree

INTRODUCTION

Manic-depressive disease (MDD; bipolar affective disorder) and alcoholism are common and severe psychiatric syndromes. Twin, adoption, and family studies are compatible with the presence of important genetic factors contributing to the etiology of a significant proportion of cases of each disorder [Kaij, 1960; Gershon et al., 1987; Blehar et al., 1988; Merikangas, 1990]. An important hypothesis for the genesis of affective disorders has been the catecholamine theory of depression [Bunney and Davis, 1965; Schildkraut, 1965]. Similarly, the effects of alcohol on reward and reinforcement mechanisms and on dopamine release suggest the involvement of dopaminergic pathways in the etiology of alcoholism [Koob, 1992; Koob and Weiss, 1992].

Due to the pivotal role of the enzyme tyrosine hydroxylase (TH) in the control of synthesis of dopamine and other catecholamines, a number of genetic linkage and relative risk association studies have been reported for this candidate gene locus, some with positive [Egeland et al., 1987; LeBoyer et al., 1990; Meloni et al., 1995] and others with negative [Kelsoe et al., 1989; Todd and O'Malley, 1989; Nöthen et al., 1990; Gill et al., 1991; Todd et al., 1996] results. Unreplicated and inconsistent results have plagued the history of linkage studies for MDD and remain the status quo, with three recent reports [Blackwood et al., 1996; Freimer et al., 1996; Ginns et al., 1996] implicating loci on five different chromosomes to bring the total number of regions showing evidence for linkage to fourteen [Risch and Botstein, 1996]. Risch and Botstein [1996] suggest that the complex nature of the genetic mechanisms underlying the disease leads to reduced power and consequent difficulty in distinguishing true from false positive results. They suggest that linkage in a small subset of families may indeed represent reality. Given the likely heterogeneous nature of both disorders studied here, we sought to apply a method of analysis that increased the power to detect an association involving only a small fraction of the disease groups with a candidate locus (TH) previously studied using less sensitive methods [Todd and O'Malley, 1989; Todd et al., 1996].

Testing haplotypes comprised of several markers at a locus provides increased power over testing individual markers. Power can be further increased by the application of cladistic analysis, an approach from population genetics which makes use of the mutational relationships among haplotypes at a locus to construct a cladogram (gene tree), which represents its evolutionary history. The cladogram, which interrelates haplotypes at a candidate locus, can be used to define the statistical analysis to test for association between affection status and increasingly inclusive branching levels of the tree. Within the cladogram, disease mutation-bearing haplotype(s) will cluster on a branch. This approach obviates testing each haplotype separately

Contract grant sponsor: National Institutes of Health; Contract grant numbers: MH31302 and DA07261.

*Correspondence to: Richard D. Todd, Ph.D., M.D., Department of Psychiatry, Campus Box 8134, Washington University School of Medicine, 4940 Children's Place, St. Louis, MO 63110.

Received 12 August 1996; Revised 21 November 1996

for association with disease status in favor of a nested contingency analysis. It replaces the subjective selection of haplotype(s) to test with the imposed logic of the evolutionary history of the locus. The implementation of this approach and its application to case/control sampling designs has been described by Templeton and coworkers [Templeton et al., 1987, 1988, 1992; Crandall, 1993; Crandall and Templeton, 1993; Templeton and Sing, 1993; Templeton, 1995]. In the present study this method was applied to the tyrosine hydroxylase locus to test for the association of manic-depressive illness and alcoholism with TH haplotype(s).

MATERIALS AND METHODS

Subjects

The participants in this study include the probands of multiple incidence manic-depressive disorder (MDD) pedigrees, probands of multiple incidence alcoholism pedigrees, and psychiatrically normal individuals. The identification, ascertainment, and diagnostic evaluation of these groups have been detailed in previous publications [Parsian et al., 1995; Suarez et al., 1994]. All participants were non-hispanic white but of diverse ethnic origin. The MDD group consisted of 81 probands (mean age 38 years; 65 percent female; 76 BP-I, 5 BP-II) who met Research Diagnostic Criteria and DSM-III-R diagnostic criteria for bipolar affective disorder. Thirty probands were from the Johns Hopkins Study of Bipolar Disorder, 14 probands were from the St. Louis site of the Collaborative Program on the Psychobiology of Depression, and 37 probands were from the St. Louis site of the NIMH Molecular Genetics Initiative for Bipolar Affective Disorder. The alcoholic group (ALC) consisted of 113 probands (mean age 40 years; 20 percent female) who met Modified Feighner Criteria for alcoholism. The psychiatrically normal control group (NL) was recruited from the St. Louis site of the Epidemiological Catchment Area (ECA) study and consisted of 80 older unrelated individuals (mean age 49 years; 49 percent female) who met no DSM-III-R criteria for, and reported no first-degree relatives with a history of affective disorders, alcoholism, psychosis, or drug use.

RFLP Genotyping

Samples of genomic DNA (250 ng–1 μg) were digested with 5–10 U/μg of restriction enzyme (RE), electrophoresed in agarose gels, and transferred at neutral pH to Hybond-N (Amersham) membrane. Membranes were hybridized with 2–4 ng/ml of ^{32}P-labeled probe (pHins 118 for *Rsa*I, *Taq*I, and *Dra*I RFLPs; pHGTH44 for *Pst*I RFLP), according to an established protocol [Lobos et al., 1989], and exposed to X-ray film at −80°C for 1–7 days.

Haplotype Estimation

Study subjects were scored for four RFLPs spanning 9 kb in the TH gene region: the polymorphic *Rsa*I, *Taq*I [Elbein et al., 1985], and *Dra*I [Sten-Linder et al., 1989] sites are within the region 2.5-0.5 kb 5′ of the first exon, and the *Pst*I site [Kelsoe et al., 1988] is located about 0.6 kb 5′ of the last intron-exon junction. The locations of the polymorphic sites were mapped using double RE digests against a background map of RE sites for relatively infrequent cutters. Four-site haplotype frequencies were estimated for each of the three groups using the algorithm of Maclean and Morton [1985] and were used to convert genotypes to haplotype numbers. Individuals with rare 4-site genotypes (more likely to be mistypings) were re-analyzed, along with their parents, if available. Two of the 25 individuals re-analyzed were deleted due to detected non-inheritance of haplotypes.

Cladogram Construction

The rules of Templeton et al. [1992] were used to construct the maximum parsimony cladogram, which interrelates the four-site haplotypes by single mutational steps. Loops in the cladogram (I–IV) indicate the presence of equally parsimonious alternative haplotype connections; this ambiguity results from homoplasy (recurrent mutation of the RE sites studied) or recombination within the region. Probabilities can be assigned to each of the alternative cladograms as described by Crandall and Templeton [1993]. Cladograms which have relatively more RE site losses than gains, and which position low-frequency haplotypes on branch tips rather than at interior locations, are predicted to be more likely. Details of the assignment of cladogram probabilities have been presented elsewhere [Crandall, 1993; Lobos, 1994].

Nested Analysis

The nested design used for statistical analysis was derived by following the rules of Templeton and coworkers [Templeton et al., 1987; Templeton and Sing, 1993]. Evolutionarily closely related haplotypes are grouped together into higher "*clades*" to test for associations between phenotype (disease status) and larger branches of the cladogram than single haplotypes. The nested design was superimposed on the most probable cladograms under two paradigms for the evolution of the TH haplotypes, homoplasy or recombination. The clades were constructed and the nested analysis done separately for each cladogram in the probable set.

Associations between haplotypes and phenotype were investigated by a series of three (the normal control and two disease groups) × n_i contingency analyses, where n_i is the number of clades in nesting category i [Templeton, 1995]. The contingency table contrasts the number of times a particular clade (haplotype or grouped haplotypes) is found in the control and case populations. A series of nested contingency tests are asymptotically independent of one another, and different contingency tests at the same clade level use nonoverlapping subsets of the data [Prum et al., 1990; Templeton, 1995]. An exact permutational chi-square test was done using the Monte-Carlo simulation-based algorithm of Roff and Bentzen [1989] as implemented by George Carmondy [personal communication]. *P*-values and standard errors were estimated from 1000 simulations for first order testing. For *P*-values near or

less than 0.05, ten thousand simulations were completed to obtain more accurate error estimates.

When a significant association was detected for the test of the three groups, additional contingency tests were done comparing each disease group with the normal control group. When a significant finding occurred within a nesting category that contained more than one evolutionary transition, 2 × 2 contingency tests were done to locate the transition showing the significant phenotypic difference. The methodology dictates that these comparisons be made only between the most closely related pairs of clades within the category.

RESULTS

Cladogram and Nesting Structure

The combined manic-depressive, alcoholic, and normal control samples included 274 genotypes divided into 26 four-site classes. For 15 of these genotype classes (44% of individuals), haplotypes were unambiguous. A minimum of nine haplotypes were required to account for all of the genotypes; the cladistic analysis was confined to that parsimonious set, using frequencies estimated [Maclean and Morton, 1985] from the combined sample. The haplotypes, and the estimated number of observations of each haplotype in the sampled populations, are presented in Table I.

The cladogram constructed from the nine 4-site haplotypes is presented in Figure 1a. The multiple closed loops (I–IV) indicate that more than one structure is consistent with the data. There are four ways to break each loop in the cladogram; doing so in all possible combinations among loops results in 185 equally parsimonious (i.e., with single mutational steps) ways to uniquely interrelate the haplotypes. The methods of Crandall and Templeton [1993; Crandall, 1993] were used to assign probabilities to the alternative cladograms, as described under Materials and Methods. The results for the 19 most probable cladograms, with a cumulative probability of 95%, are presented in Figure 2. The first four cladograms represent about 80%, the remaining 15 cladograms each less than 3%, of the total probability. Therefore, the contingency analysis was confined to the first four cladograms (Fig. 3a).

TABLE I. Observed and Estimated TH Haplotypes for Three Study Groups

		MDD		ALC		NL	
Haplotype[a]	R T D P[b]	obs[c]	est[d]	obs	est	obs	est
1	−+++	19	27.5	18	24.9	16	22.5
2	++++	1	1.4	1	1.6	0	0.0
3	++−+	30	37.1	40	52.5	22	37.5
4	+−−+	9	17.0	20	36.0	13	20.9
5	+−−−	1	1.0	0	0.0	4	6.1
6	++−−	20	42.8	28	56.5	15	29.4
7	+++−	6	7.7	2	2.4	1	3.0
8	−++−	5	26.4	18	51.1	16	37.4
9	−+−−	1	1.1	1	1.0	3	3.0

[a]While two additional haptotypes were assigned small, but non-zero, frequencies by the haplotype estimation algorithm, all genotypes can be accounted for with only nine haplotypes.
[b]Allele status for the restriction enzyme sites as they occur 5′ to 3′ in the TH gene: R, *Rsa*I; T, *Taq*I; D, *Dra*I; *P*, *Pst*I; [+] site present, [−] site absent.
[c]Haplotypes directly observed in unambiguous genotype classes; includes haplotypes from genotype classes that were unambiguous only when constrained by nine allowed haplotypes.
[d]Haplotype frequencies were estimated for each group separately by the Maclean and Morton [1985] algorithm; they were multiplied by the number of chromosomes scored.

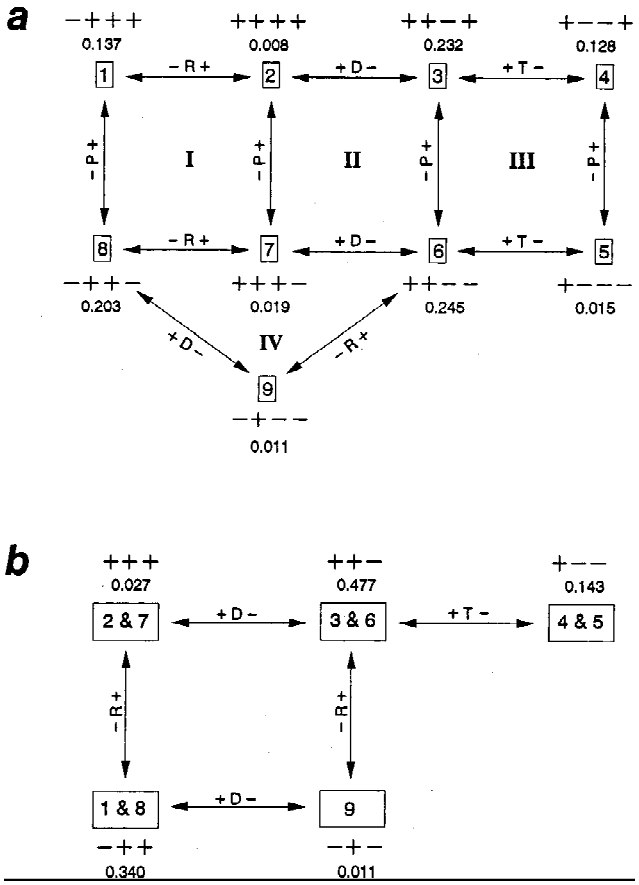

Fig. 1. **(a)** The cladogram set interrelating the nine TH haplotypes, as estimated by the Templeton et al. [1992] algorithm. The frequency of each haplotype (estimated from the combined groups) is indicated below the allelic state (+ or −) for each of the four restriction enzyme (RE) site polymorphisms, listed 5′—3′, as in Table 1. Each double-headed arrow indicates a mutational event at one of the polymorphic sites. The allelic state (presence [+], absence [−]) for the haplotype at each end is indicated on either side of the RE site designated adjacent to the arrow (R = *Rsa*I, T = *Taq*I, D = *Dra*I, P = *Pst*I). The multiple loops (I–IV) indicate ambiguity in the evolutionary history of the locus and may result from homoplasy (multiple independent mutations leading to the same allelic state) or recombination. **(b)** The cladogram set estimated for the three RFLPs of the 5′ region under the rules of Templeton and Sing [1993], allowing for recombination between the *Dra*I and *Pst*I sites. Haplotypes are boxed together which are identical for the three sites in the 5′ region. The 3′ *Pst*I site is analyzed separately.

Eight of the nine haplotypes define pairs that are identical for the three 5′ RFLPs of TH (*Rsa*I, *Taq*I, *Dra*I) but different at the 3′ *Pst*I site. The two most likely explanations for this state are (1) frequent recombination in the intervening region, and (2) recurrent independent mutations creating or eliminating the *Pst*I site (homoplasy). The *Pst*I site is separated from the other sites by a greater distance (7–9 kb), making recombination plausible. In that case, the rules of Templeton and Sing [1993] suggest analyzing the 5′ and 3′ regions separately, resulting in the cladogram for the 5′ region shown in Figure 1b. The 3′ region then

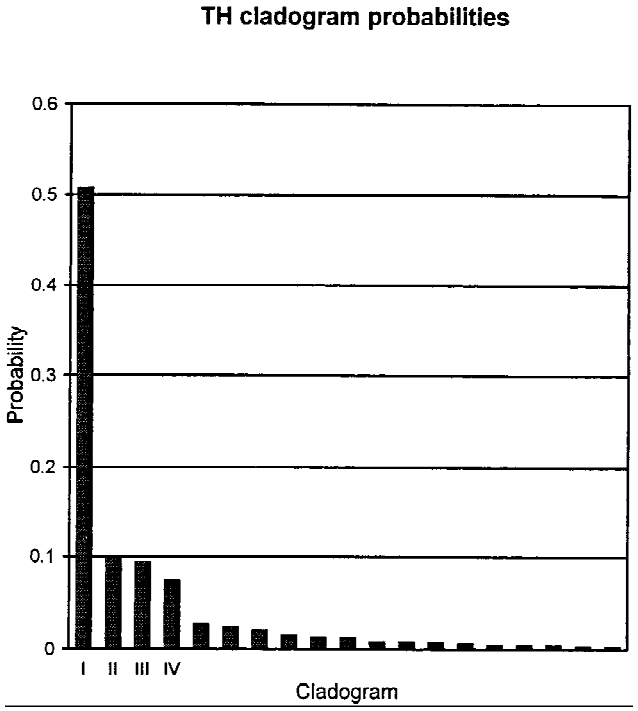

Fig. 2. Probabilities assigned by the methods of Crandall and Templeton [1993; also see Crandall, 1993 and Lobos, 1994] to the 95% probability subset of the 185 alternative cladograms that result from breaking all loops of the structure in Figure 1a in all possible ways. A single cladogram is allocated 51% of the probability; the four most likely cladograms represent a 78% probability set, and are labeled I–IV as in Figure 3a.

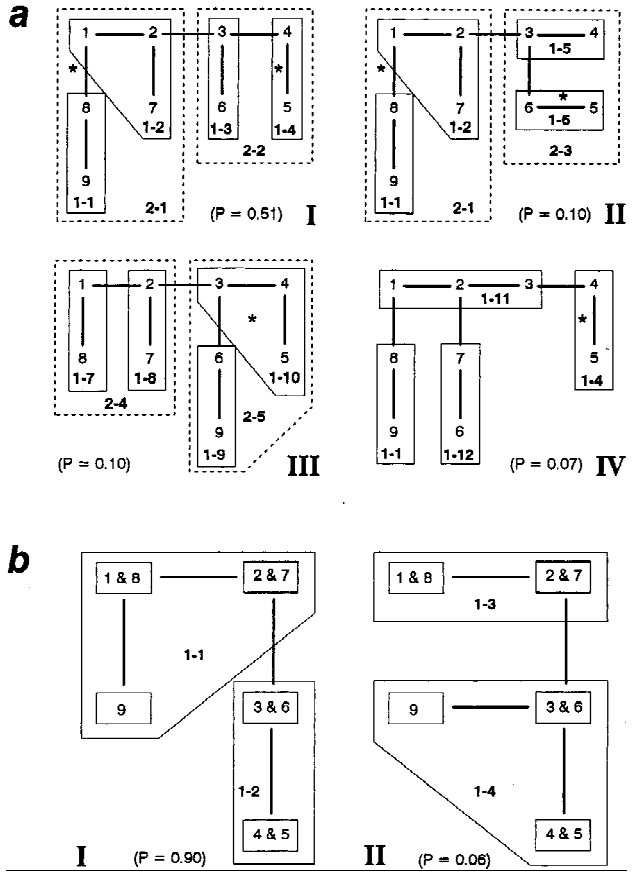

Fig. 3. The nested statistical design superimposed on the most probable cladograms by the rules of Templeton and coworkers [Templeton et al., 1987; Templeton and Sing, 1993]. The *P* value is the probability assigned to the cladogram according to the methods of Crandall and Templeton (see text). The evolutionary transitions associated with significant group differences are indicated by asterisks. Increasingly larger branches (clades) of the gene tree (cladogram) are tested in a nested fashion for association with the disease phenotypes. The final level of nesting always includes the entire cladogram. **(a)** The four most probable cladograms from Figure 1a (see Fig. 2). Haplotypes enclosed by solid lines indicate 1-step clades (designated by 1-#). Dashed lines enclose 2-step clades (designated by 2-#), except for cladogram IV where that level of nesting includes the entire cladogram. Each clade that is unique over all four cladograms is numbered sequentially. **(b)** The nested design for the two cladograms included in the 95% probable set for the 5′ region of the TH gene, showing two 1-step clades for each cladogram. The two-step clades include the entire cladogram.

consists solely of the *Pst*I site. The contingency analysis was confined to two of the four alternative cladograms for the 5′ region, which represent 96% of the total probability (Fig. 3b).

Nested Contingency Analysis

Since it can not be determined from our data whether cladogram ambiguity is due to homoplasy or recombination, the data were analyzed under both models. The nested designs used for the contingency analyses of the most probable cladograms under both paradigms are illustrated in Figure 3a,b. The data used are given in Table I, with estimated haplotype numbers rounded to the nearest integer. The results for the nested three-group analyses are provided in Table II.

When recombination is allowed for (Fig. 3b), and the cladograms for 3-site haplotypes are analyzed, there is no significant evidence for heterogeneity among the groups. The 3′ *Pst*I RFLP, analyzed separately, similarly shows no significant allele frequency differences among the three groups. For comparison, a standard contingency analysis of the five 3-site haplotypes also shows no significant difference among the three groups (Table II). However, a similar analysis of the nine 4-site haplotypes showed statistical evidence for frequency differences across the three groups ($P = 0.01$).

There are significant permutational chi-squares associated with four comparisons within the homoplasy cladograms (Fig. 3a): within the 2-step clade 2-1, and within the 1-step clades 1-4, 1-6, and 1-10. The evolutionary transitions that are associated with each of the significant differences, as identified by these comparisons, are indicated by asterisks in Figure 3a. While each of the first three chi-squares is associated with a single evolutionary transition, there are two transitions within clade 1–10: the mutations that connect haplotype 4 with both haplotype 3 and haplotype 5. Examination of Figure 3a and Table II indicate that each of these transitions has already been tested: the former within clade 1–5 and the latter within clade 1–4. The results show that the transition between haplotypes 4 and 5 is associated with the significant group difference.

Further testing was done to determine whether these significant frequency differences were associated with one or the other of the disease groups (Table III). The chi-square of modest significance (clade 2-1, $P =$

TABLE II. Results of Nested Contingency Analyses of Data for Three Study Groups*

	Chi-square statistic	Probability	S.E.
Homoplasy cladograms			
Zero-step clades			
within 1–1	1.77	0.43	0.02
within 1–2	5.09	0.27	0.01
within 1–3	1.80	0.41	0.02
within 1–4	9.93	0.0075	0.0009
within 1–5	1.20	0.55	0.02
within 1–6	13.57	0.0009	0.0003
within 1–7	4.95	0.085	0.003
within 1–8	3.60	0.22	0.01
within 1–9	3.75	0.16	0.01
within 1–10	11.66	0.017	0.001
within 1–11	3.36	0.54	0.02
within 1–12	4.89	0.084	0.003
One-step clades			
within 2–1	7.96	0.023	0.002
within 2–2	2.91	0.24	0.01
within 2–3	1.12	0.61	0.02
within 2–4	5.31	0.083	0.003
within 2–5	2.40	0.31	0.01
within cladogram IV	9.43	0.14	0.01
Two-step clades			
within cladograms I and II	1.26	0.54	0.02
within cladogram III	0.80	0.68	0.01
Recombination cladograms			
Zero-step clades			
within 1–1	7.39	0.11	0.02
within 1–2	2.91	0.24	0.02
within 1–3	5.31	0.083	0.003
within 1–4	5.44	0.23	0.01
One-step clades			
within cladogram I	1.26	0.54	0.02
within cladogram II	0.80	0.68	0.01
*Pst*I site	0.01	1.00	0.00
Nine haplotypes	31.18	0.010	0.001
Five haplotypes	11.29	0.19	0.01

*The nested design is shown in Figure 3. Zero-step clades are the individual haplotypes. The result for a given clade is listed only once even though it may occur in more than one cladogram. A contingency chi-square was calculated, and its significance was determined by 1,000 random permutations that preserved the marginal values. The probability refers to the frequency with which these randomly generated chi-square statistics were equal to or greater than the observed chi-square. For *P*-values near or less than 0.05, 10,000 random permutations were used.

TABLE III. Results of Analyses to Determine the Source of Significant Group Differences

	Chi-square statistic	Probability	S.E.
Within clade 2–1			
MDD *vs.* normal	4.84	0.034	0.002
ALC *vs.* normal	0.11	0.863	0.003
Within clade 1–4			
MDD *vs.* normal	2.28	0.218	0.004
ALC *vs.* normal	8.84	0.0049	0.0007
Within clade 1–6			
MDD *vs.* normal	5.14	0.044	0.002
ALC *vs.* normal	9.98	0.0032	0.0006
Within clade 1–10			
MDD *vs.* normal	3.19	0.227	0.004
ALC *vs.* normal	8.90	0.0088	0.0009
Nine haplotypes			
MDD *vs.* normal	13.22	0.089	0.003
ALC *vs.* normal	15.05	0.047	0.002

0.023; detected in cladograms I and II, Fig. 3a) associated with the transition between clades 1-1 and 1-2 was due to differences between the MDD and normal control groups. In contrast, significant differences between the alcoholic and normal control groups were associated with the transitions within the 1-step clades 1–4, 1–6, and 1–10. In all four cladograms, the significant difference involved the connection of the 4-site haplotype 5 to the rest of the cladogram. It should be noted that haplotype 5 is absent in the alcoholic group, rather than at higher frequency. The standard contingency analysis of the nine 4-site haplotypes showed marginally significant statistical evidence for frequency differences between the alcoholic and normal control groups ($P = 0.047$; Table III).

DISCUSSION

The current study was designed to further investigate tyrosine hydroxylase as a candidate gene for manic-depressive illness and alcoholism by using information from several polymorphic markers, and employing an evolutionary context to guide statistical analysis. A cladogram was constructed for TH that interrelated the estimated haplotypes with the minimum number of mutational changes, in order to provide a design for a nested contingency analysis. This approach, which has been useful in the quantitative analysis of alcohol dehydrogenase gene activity [Templeton et al., 1987] and in the qualitative analysis of the apoE locus and Alzheimers disease [Templeton, 1995], makes maximal use of available genotype and phenotype information. If present, a disease mutation should be embedded within the evolutionary tree of genetic variation at the locus; thus, the associated haplotype(s) should cluster on some branch of the tree. This approach provides the statistical power to localize, within a cladogram of haplotypes, multiple mutations of small effect or mutations which affect only a fraction of the individuals with a given phenotype.

Previous findings of association between manic-depressive illness and the tyrosine hydroxylase gene were reported by LeBoyer et al. [1990] for *Taq*I and *Bgl*II RFLP alleles and by Meloni et al. [1995] for a particular genotype of the intron I tetranucleotide repeat polymorphism. For essentially the same population as used in this study, we have previously reported no evidence of association between any of these three polymorphisms and MDD [Todd et al., 1996]. In the present study, we have extended these findings by comparing TH haplotype frequencies for the MDD and control groups within an evolutionary context, using methodology better able to detect association with a subset of the disease group.

Of the aforementioned polymorphisms, the *Taq*I RFLP is included here in the haplotypes constructed for the TH locus. The *Bgl*II RFLP is due to a hypervariable repeat region (HVR) that LeBoyer et al. [1990] scored as two alleles. The situation is considerably more complex than that [McGinnis and Spielman, 1994], and it is not clear that analysis of the HVR alleles as two size classes is genetically meaningful. The relative probability of transitions among the different size classes of the HVR is unknown; therefore, it is not

amenable to a probabilistic treatment of alternative cladograms and has not been included in this analysis. The same considerations apply to the tetranucleotide repeat polymorphism.

Using a cladistics-based contingency analysis, we have detected nominally significant associations of genetic variation at the TH gene with both MDD and alcoholism. The association with MDD (within clade 2-1) was of modest significance, was detected with two of the four most probable cladograms, and did not involve a change at the *Taq*I site. A marginally significant ($P = 0.044$) association with MDD was also found within clade 1–6 for one cladogram. The association with alcoholism involved the transition between haplotype 5 and either haplotype 4 or 6, and was due to the absence of haplotype 5 in the ALC group. Although found with all four cladograms, such a "protective" effect, if real, would be of small importance, since haplotype 5 is also rare in normal controls.

These findings of nomimally significant disease association should be viewed with caution. First, the ambiguity of the cladogram required contingency analyses using several alternative nested designs. Signifigant group differences were detected between clades that were not defined in the same way for all probable cladograms, hence were cladogram-dependent. Certainly, the nominal *P*-values should be interpreted with the knowledge that they were obtained under conditions of multiple, non-independent tests, for which the appropriate *P*-value is uncertain. Second, haplotype numbers were estimated rather than directly observed. Analysis of haplotypes, rather than single polymorphisms, requires family data to determine phase or relies on estimation algorithms to determine haplotypes. While the estimation of haplotype frequencies adds another source of uncertainty, any population sample provides only an estimate of the population characteristics. Perhaps of greater concern, ethnic heterogeneity among case and control samples may create the appearance of disease association. Efforts are underway to determine whether analysis of parent-offspring trio data for this locus, using nontransmitted haplotypes as the control sample, will replicate the results of this study.

Given these concerns, it is appropriate to examine the impact on the study results of minor changes in the estimated haplotype frequencies. The proportion of the haplotypes that were directly observed varies (19—100% observed, Table I) among haplotypes and diagnostic groups. When a rare haplotype is not unambiguously observed, resulting in a negligible frequency estimate, the cell with zero observations may have a large impact on the contingency analysis. Therefore, we explored the consequences of adding one individual to the cell of the contingency tables with the lowest value, while retaining the marginal sums. This resulted in non-significant values for the contingency analyses in all cases except clade 1–6 for alcoholics versus normals, which was then of marginal significance ($P = 0.034$). Consequently, we suggest that the observed group differences be viewed skeptically unless replicated.

When cladistic analysis results in a single cladogram, it provides an evolutionary design for a nested statistical analysis that groups the more closely related haplotypes to test for association with disease at higher levels in the hierarchy. In doing so, it (1) provides increased power to detect association with a small subset of the disease group or association with several closely related haplotypes, (2) removes unintentional bias inherent in "previewing" the haplotype distributions and selecting specific haplotypes to test for association, (3) restricts the contingency analyses to the evolutionarily meaningful nested set, (4) circumvents the problem of multiple tests when analyzing multiple alleles (haplotypes) at a locus, and (5) indicates which haplotypes merit further investigation for elucidation of the disease susceptibility mutation.

Uncertainty regarding the "true" cladogram decreases the value of this approach. However, even when the cladogram is ambiguous, it can increase power if it consistently groups the disease-associated haplotypes. In any case, the exercise of constructing a cladogram reveals important information about the locus being studied, such as whether a history of recurrent mutation or recombination is required to explain the present array of haplotypes. Such information is relevant to all methods of association or disequilibrium analysis. Even an ambiguous cladogram can shed light on evolutionary processes or reveal questions that need to be addressed. For example, all cladograms for the TH gene place frequent haplotypes (3 and 6, 1 and 8) on opposite sides of rare haplotypes (2, 7, or 9). This might seem evolutionarily implausible. However, this state has been found for other loci as well [Lobos, 1994]. If genetic drift has been an important factor in the history of human populations, producing large and relatively rapid fluctuations in allele (haplotype) frequencies, this state may not be an unusual occurrence.

In the face of ambiguity, the rules suggested by Templeton and colleagues [Crandall and Templeton, 1993; Crandall, 1993] for estimating cladogram probabilities can greatly reduce the number of plausible alternatives. However, it is still uncertain whether those rules are generally appropriate for selecting the "true" cladogram or, at least, reducing the alternatives to a small number which include the "true" cladogram. Other rules useful for cladogram estimation or testing may remain to be discovered. It may prove necessary to use more polymorphic sites than studied here, or to use more closely spaced sites, or to sequence several individuals from each haplotype, in order to determine whether recombination or recurrent mutation has obscured the relationships among haplotypes. Investigation of a variety of candidate genes in humans will be required to provide answers to these questions. The plethora of conflicting reports of association of single markers at many loci with many diseases suggests that effort invested in finding an alternative approach would be well spent.

ACKNOWLEDGMENTS

We thank Alan Templeton for advice and Brian Suarez for reading an earlier version of this report. We thank an anonymous reviewer for helpful comments. We thank Janice Hickok for technical assistance and

review of data. This work was supported by grants MH31302 and DA07261 of the National Institutes of Health.

REFERENCES

Blackwood DHR, He L, Morris SW, McLean A, Whitton C, Thomson M, Walker MT, Woodburn K, Sharp CM, Wright AF, Shibasaki Y, St. Clair DM, Porteous DJ, Muir WJ (1996): A locus for bipolar affective disorder on chromosome 4p. Nat Genet 12:427–430.

Blehar MC, Weissman MM, Gershon ES, Hirschfeld RM (1988): Family and genetic studies of affective disorders. Arch Gen Psychiatry 45:289–292.

Bunney WE Jr., Davis JM (1965): Norepinephrine in depressive reactions. Arch Gen Psychiatry 13:483–494.

Crandall KA (1993): "Molecular Systematics and Evolutionary Biology of the Crayfish Subgenus *Procericambarus.*" St. Louis: Washington University (Ph.D. Thesis).

Crandall KA, Templeton AR (1993): Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. Genetics 134:959–969.

Egeland JA, Gerhard DS, Pauls DL, Sussex JN, Kidd KK, Allen CR, Hostetter AM, Housman DE (1987): Bipolar affective disorder linked to DNA markers on chromosome 11. Nature 325:783–787.

Elbein SC, Corsetti L, Permutt MA (1985): New polymorphisms at the insulin locus increase its usefulness as a genetic marker. Diabetes 34: 1139–1144.

Freimer NB, Reus VI, Escamilla MA, McInnes LA, Spesny M, Leon P, Service SK, Smith LB, Silva S, Rojas E, Gallegos A, Meza L, Fournier E, Baharloss S, Blankenship K, Tyler DJ, Batki S, Vinogradov S, Weissenbach J, Barondes SH, Sandkuijl LA (1996): Genetic mapping using haplotype, association and linkage methods suggests a locus for severe bipolar disorder (BPI) at 18q22–q23. Nat Genet 12:436–441.

Gershon ES, Berrettini W, Nurnberger J Jr, Goldin LR (1987): Genetics of affective illness. In Melzer HY (ed): "Psychopharmacology." New York: Raven Press, pp. 481–493.

Gill M, Castle D, Hunt N, Clements A, Sham P, Murray RM (1991): Tyrosine hydroxylase polymorphisms and bipolar affective disorder. J Psychiatr Res 25:179–184.

Ginns EI, Ott J, Egeland JA, Allen CR, Fann CSJ, Pauls DL, Weissenbach J, Carulli JP, Falls KM, Keith TP, Paul SM (1996): A genome-wide search for chromosomal loci linked to bipolar affective disorder in the Old Order Amish. Nat Genet 12:431–435.

Kaij L (1960): "Alcoholism in twins: studies on the etiology and sequels of abuse of alcohol." Stockholm: Almqvist and Wiksell.

Kelsoe JR, Ginns EI, Egeland JA, Gerhard DS, Goldstein A, Bale SJ, Pauls DL, Long RT, Kidd KK, Conte G, Housman DE, Paul SM (1989): Reevaluation of the linkage relationship between chromosome 11p loci and the gene for bipolar affective disorder in the Old Order Amish. Nature 342:238–243.

Kelsoe JR, Stubblefield BK, Ginns EI (1988): Human tyrosine hydroxylase (TH) genomic frament (pHGTH4) identifies a *Pst*I polymorphism. Nucl Acids Res 16:7760.

Koob GF (1992): Drugs of abuse: anatomy, pharmacology and function of reward pathways. Trends in Pharmacol Sci 13:177–184.

Koob GF, Weiss F (1992): Neuropharmacology of cocaine and ethanol dependence. Recent Dev Alcoholism 10:201–233.

LeBoyer M, Malafosse A, Boularand S, Campion D, Gheysen F, Samolyk D, Henriksson B, Denise E, des Lauriers A, Lepine J-P, Zarifian E, Clerget-Darpoux F, Mallet J (1990): Tyrosine hydroxylase polymorphisms associated with manic-depressive illness. Lancet 335:1219.

Lobos EA (1994): "Genetic Linkage Mapping and Linkage Disequilibrium Analysis of Restriction Fragment Length Polymorphisms of the Human Muscle Nicotinic Accetylcholine Receptor Genes." St. Louis: Washington University (Ph.D. Thesis).

Lobos EA, Rudnick CH, Watson MS, Isenberg KE (1989): Linkage disequilibrium study of RFLPs detected at the human muscle nicotinic acetylcholine receptor subunit genes. Am J Hum Genet 44:522–533.

Maclean CJ, Morton NE (1985): Estimation of myriad haplotype frequencies. Genet Epidemiol 2:263–272.

McGinnis RE, Spielman RS (1994): Linkage disequilibrium in the insulin gene region: size variation at the 5′ flanking polymorphism and bimodality among "Class I" alleles. Am J Hum Genet 55:526–532.

Meloni R, LeBoyer M, Bellivier F, Barbe B, Samolyk D, Allilaire JF, Mallet J (1995): Association of manic-depressive illness with tyrosine hydroxylase microsatelite marker. Lancet 345:932.

Merikangas KR (1990): The genetic epidemiology of alcoholism. Psychol Med 20:11–25.

Nöthen M, Korner J, Lanczik M, Fritze J, Propping P (1990): Tyrosine hydroxylase polymorphisms and manic-depressive illness. Lancet 336: 575.

Parsian A, Chakraverty S, Todd RD (1995): Possible association between the dopamine D3 receptor gene and bipolar affective disorder. Am J Med Genet 60:234–237.

Prum B, Guilloud-Bataille M, Clerget-Darpoux F (1990): On the use of χ^2 tests for nested categorized data. Ann Hum Genet 54:315–320.

Risch N, Botstein D (1996): A manic depressive history. Nat Genet 12:351–353.

Roff DA, Bentzen P (1989): The statistical analysis of mitochondrial DNA polymorphisms: χ^2 and the problem of small samples. Mol Biol Evol 6:539–554.

Schildkraut JJ (1965): The catecholamine hypothesis of affective disorders: A review of supporting evidence. Am J Psych 122:509–522.

Sten-Linder M, Granqvist M, Stern I, Bell GI, Luthman H (1989): DraI and PstI RFLPs in the tyrosine hydroxylase (TH) and insulin gene (INS) region of chromosome 11. Nucl Acids Res 17:5873.

Suarez BK, Parsian A, Hampe CL, Todd RD, Reich T, Cloninger CR (1994): Linkage disequlibria at the D2 dopamine receptor locus (DRD2) in alcoholics and controls. Genomics 19:12–20.

Templeton AR (1995): A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping or DNA sequencing. V. Analysis of case/control sampling designs: Alzheimer's disease and the Apoprotein E locus. Genetics 140:403–409.

Templeton AR, Boerwinkle E, Sing CF (1987): A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. I. Basic theory and an analysis of alcohol dehydrogenase activity in *Drosophila.* Genetics 117:343–351.

Templeton AR, Crandall KA, Sing CF (1992): A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. Genetics 132:619–633.

Templeton AR, Sing CF (1993): A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. IV. Nested analysis under cladogram uncertainty and recombination. Genetics 134:659–669.

Templeton AR, Sing CF, Kessling A, Humphries S (1988): A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. II. The analysis of natural populations. Genetics 120:1145–1154.

Todd RD, Lobos EA, Parsian A, Simpson S, DePaulo JR (1996): Manic-depressive illness and tyrosine hydroxylase markers. Lancet 347:1634.

Todd RD, O'Malley KL (1989): Population frequencies of tyrosine hydroxylase restriction fragment length polymorphisms in bipolar affective disorder. Biol Psychiatry 25:626–630.